

Supplementary Information

Dynamics of metatranscription in the inflammatory bowel disease gut microbiome

Melanie Schirmer, Eric A. Franzosa, Jason Lloyd-Price, Lauren J. McIver, Randall Schwager, Tiffany W. Poon, Ashwin N. Ananthakrishnan, Elizabeth Andrews, Gildardo Barron, Kathleen Lake, Mahadev Prasad, Jenny Sauk, Betsy Stevens, Robin G. Wilson, Jonathan Braun, Lee A. Denson, Subra Kugathasan, Dermot P.B. McGovern, Hera Vlamakis, Ramnik J. Xavier, Curtis Huttenhower

Supplementary Figure 1

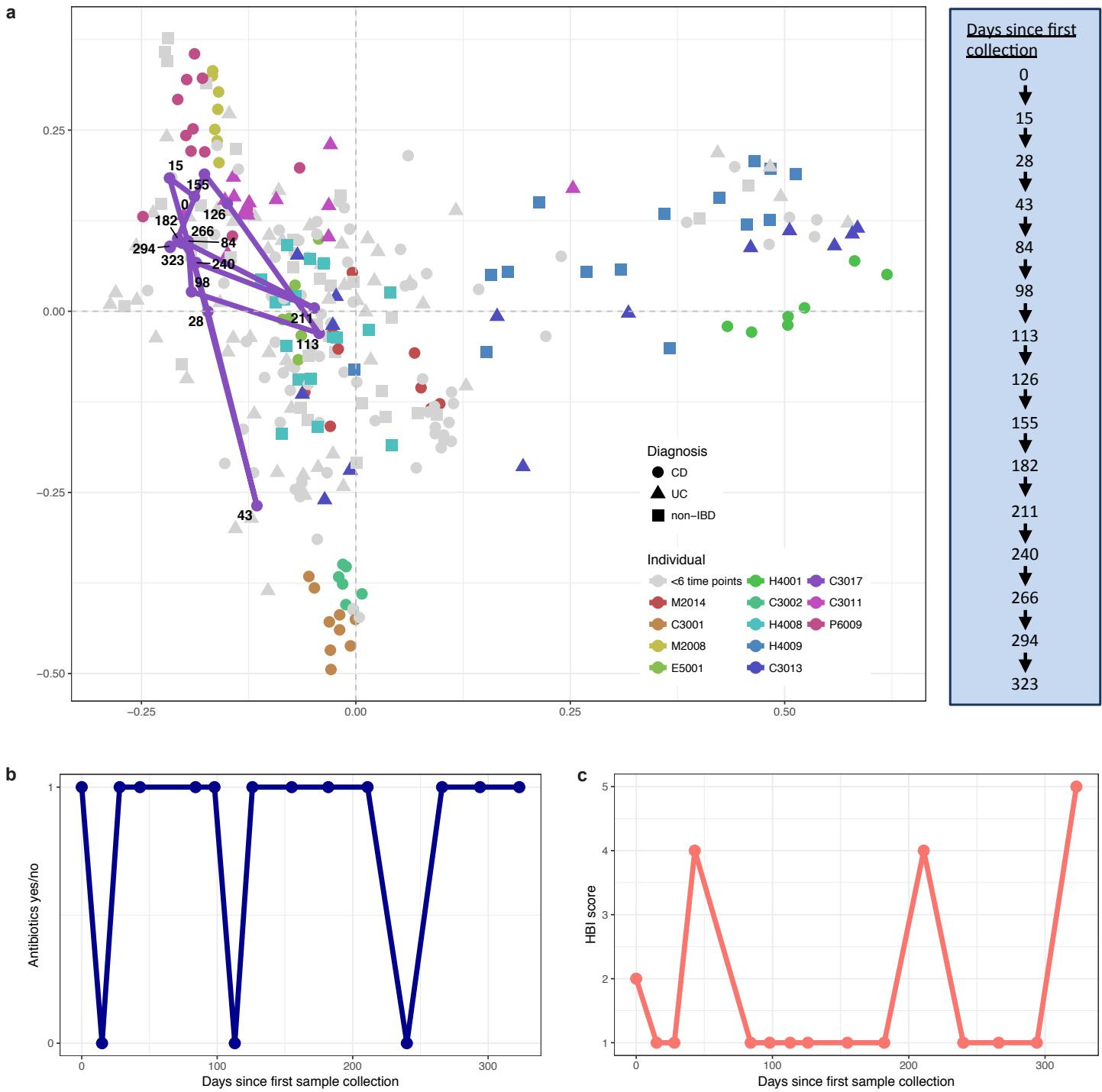


Figure S1: Longitudinal variation in a CD patient. **(a)** Same principal coordinate analysis (PCoA) as in Fig. 1b comparing species composition of the samples using Bray-Curtis distance ($n=300$). Additionally, the time course of one CD patient (C3017) is indicated, illustrating the variability in species composition over time including three outlier samples. **(b)** Antibiotic treatment (1=yes, 0=no) for patient C3017 for all time points ($n=15$). The patient received antibiotic treatment during most of the sample collection, except for 3 time points. Two of the outliers on the PCoA plot coincide with changes in antibiotic treatment (43 and 113 days after first sample collection). **(c)** Changes in disease severity for patient C3017 over time ($n=15$) as measured by the Harvey-Bradshaw Index (HBI). The third outlier sample on the PCoA, which was obtained 211 days after the first sample collection, co-occurred with a spike in the HBI score increasing by 3 points.

Supplementary Figure 2

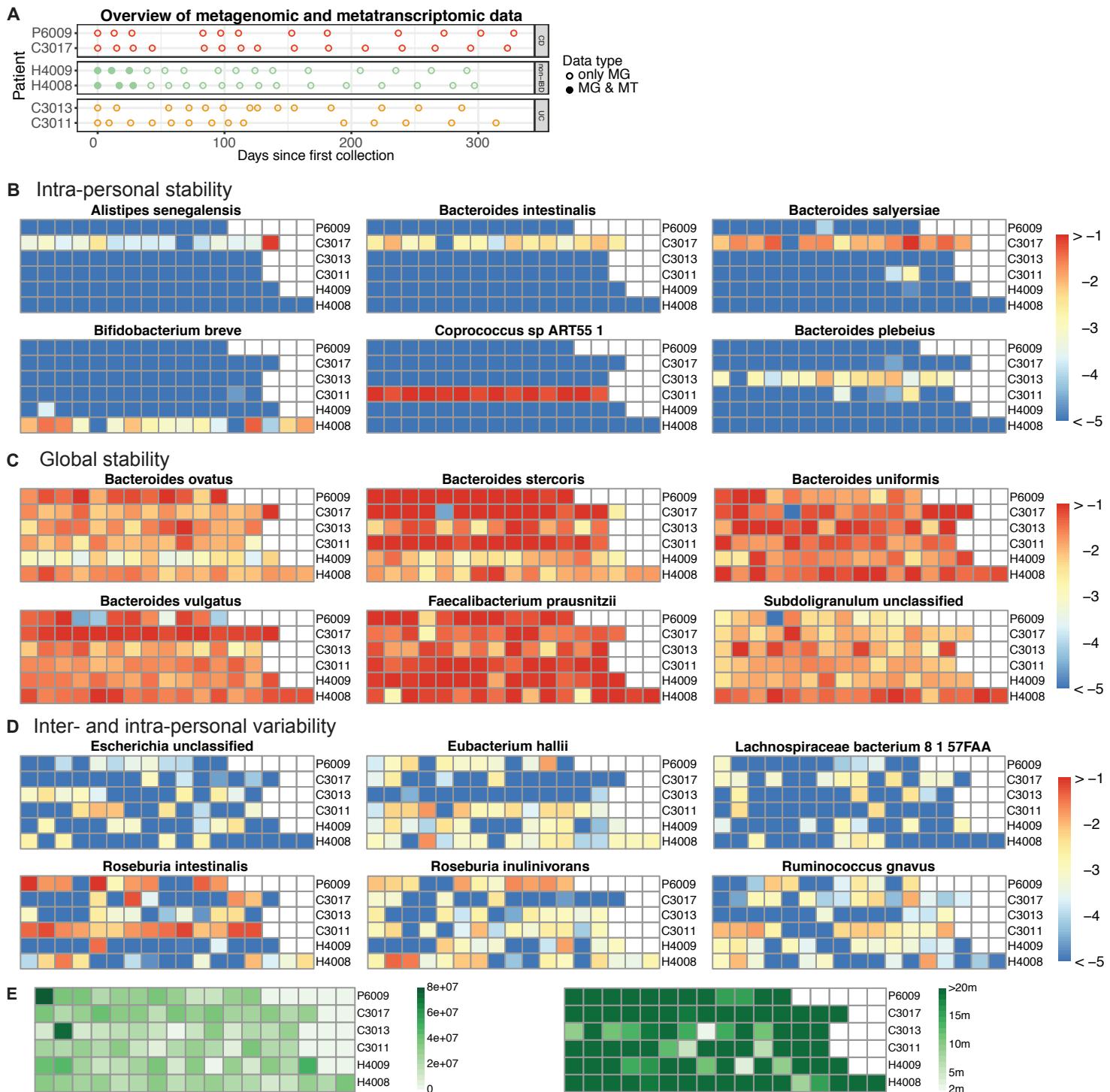


Figure S2: (a) Sampling times and sample types (metagenomics/metatranscriptomics) for all long time courses. Each row represents one patient, including 2 CD (red), 2 non-IBD controls (green) and 2 UC patients (yellow). (b-d) Inter- and intra-personal dynamic patterns of microbial species. Each row shows the relative abundance (log10 scale) of a particular species across all consecutive time points from a particular patient. The upper threshold for the relative abundances is 10%, the lower threshold for detection is 0.001% (10^{-5}). The same color scale was used for all plots. (e) Number of reads per sample. Total number of reads per sample for the long time courses (after quality control) that were used as input to infer the taxonomic profiles are shown on the left. The minimum number of reads was 2 million. The scaled version is displayed on the right with an upper threshold of 20 million reads.

Supplementary Figure 3

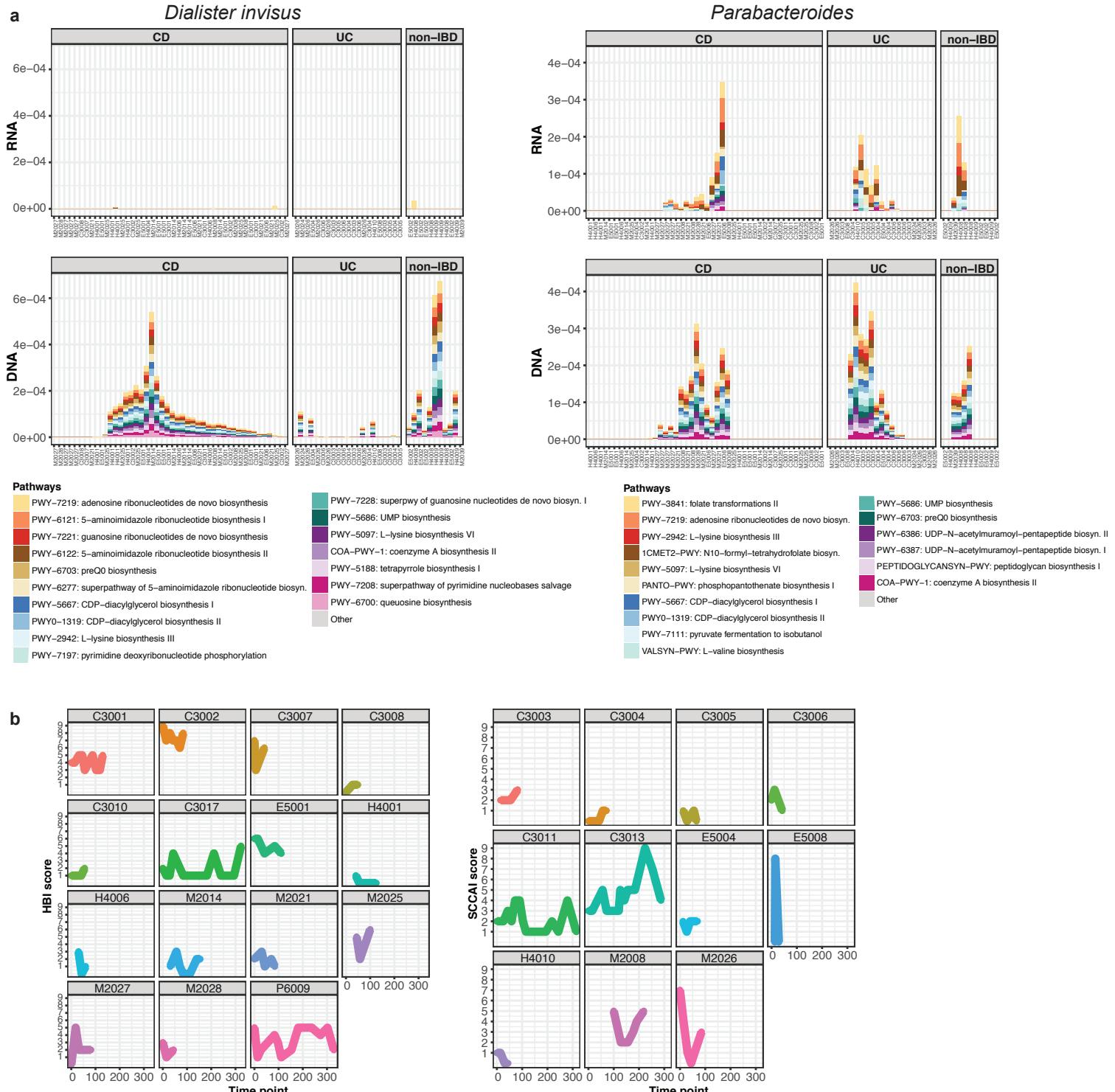


Figure S3: (a) Differences in functional potential and functional activity of microbial species. Relative contribution of a species (*Dialister invisus* on the left, *Parabacteroides merdae* on the right) to all pathways. Each column represents one sample. Samples are grouped by diagnosis and ordered based on similarity of pathway contribution (Methods). The 20 most abundant pathways in the metagenomic and metatranscriptomic data are indicated by color and the contribution to all other pathways is summarized as “Others”. (b) Distribution of disease severity scores for all patients with at least three measurements. Each panel represents one patient. The Harvey-Bradshaw Index (HB1) was measured for CD patients (left plot) and the Simple Clinical Colitis Activity Index (SCCAI) was measured for UC patients (right plot).

Supplementary Figure 4

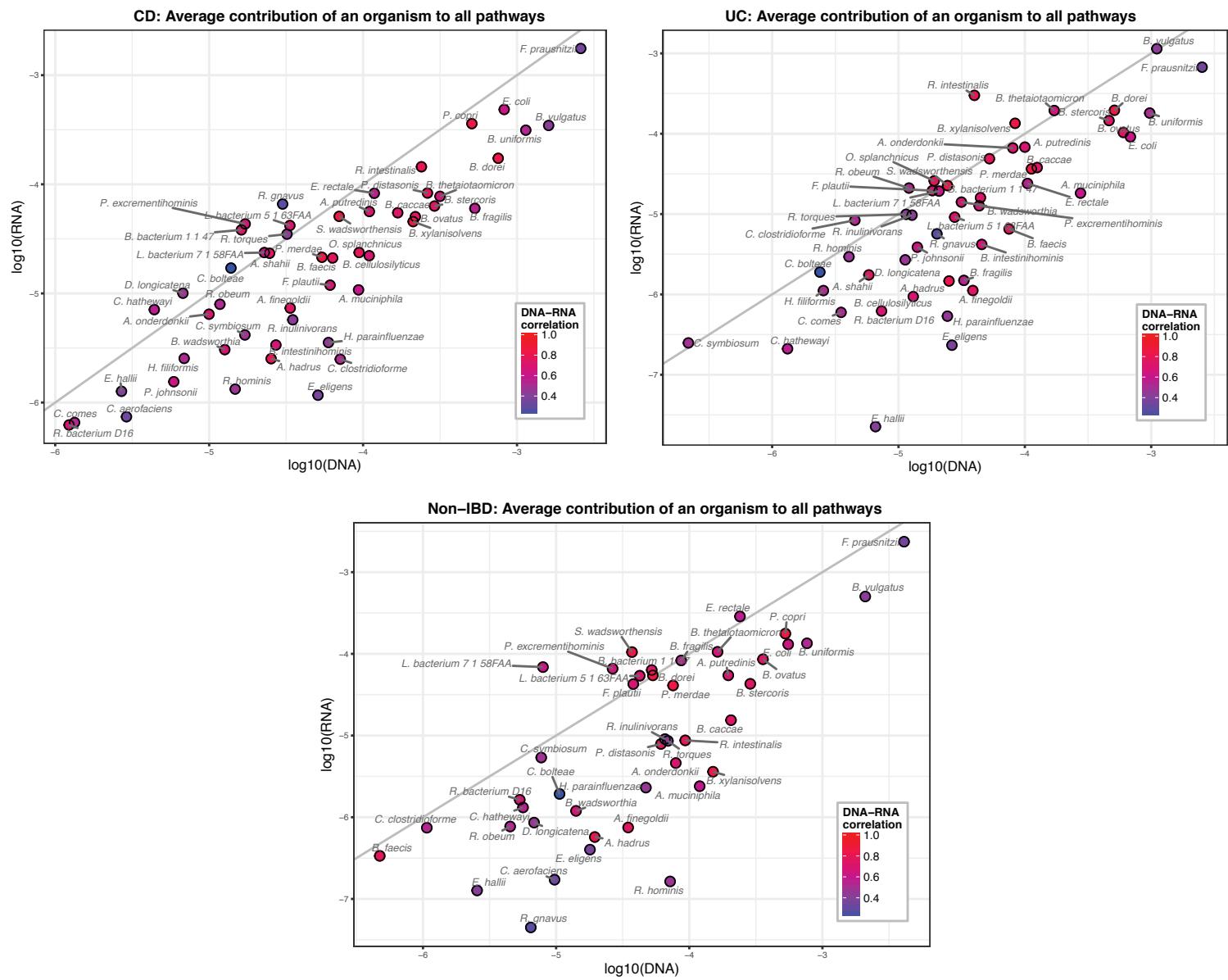
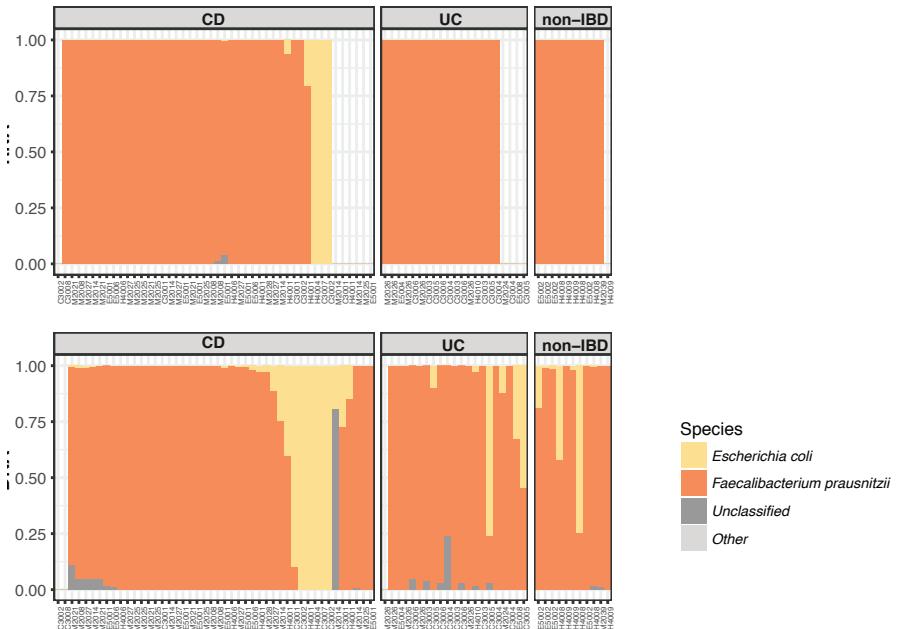


Figure S4: Average contribution of an organism to all pathways stratified by disease phenotype. Analogously to Fig. 2A, we computed the mean contribution of a microbial species (each point represents one species) to all pathways per participant and subsequently took the mean across all participants to compare DNA and RNA level contributions for each disease group (CD n=46, UC n=21, non-IBD n=11).

Supplementary Figure 5

a GALACT-GLUCUROCAT-PWY: superpathway of hexuronide & hexuronate degradation



b PWY-7219: adenosine ribonucleotides de novo biosynthesis

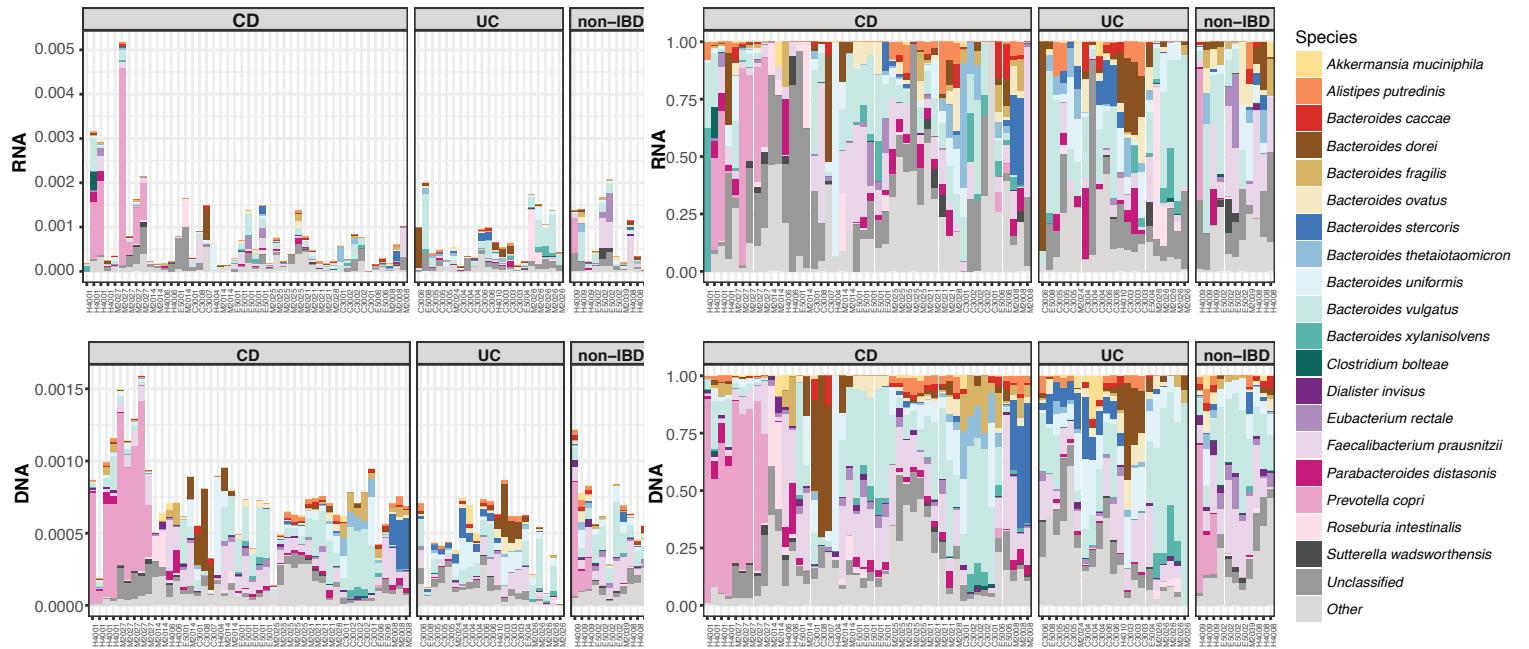
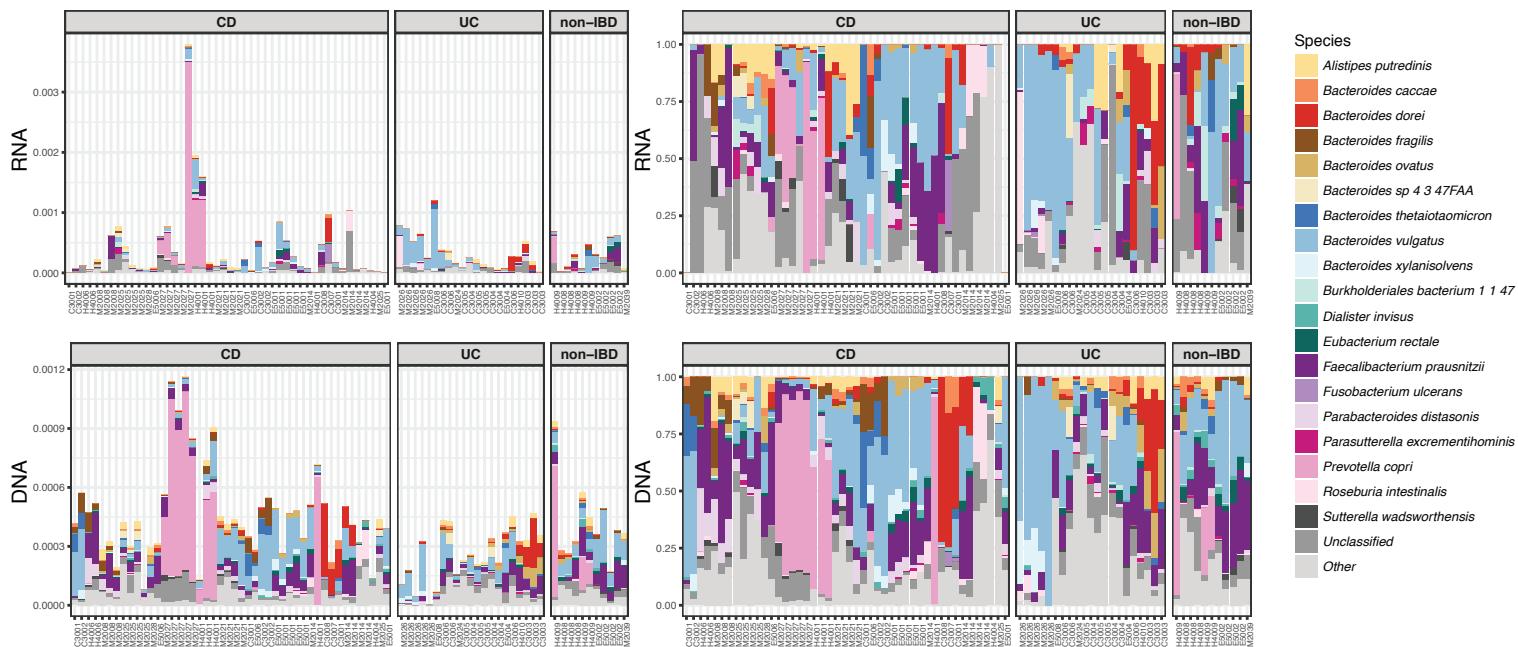


Figure S5: (a) Pathway with lowest contributonal alpha diversity on DNA and RNA level. Relative contribution of up to 20 species (ranked by average contribution) in the metagenomic and metatranscriptomic data ($n=78$) for the superpathway of hexuronide and hexuronate degradation (analogously to Fig. 3c). Each column shows the relative species-contribution to the pathway in a particular sample. Samples are stratified by disease phenotype and order based on similarity of species-contribution (Methods). (b) Pathways with the greatest contributonal alpha diversity. In addition to the scaled versions (shown on the right), the corresponding plots showing the relative abundance of the pathway in the samples ($n=78$) are displayed on the left.

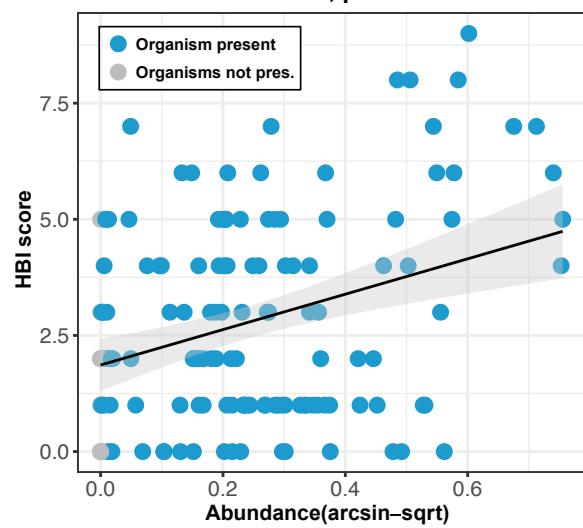
Supplementary Figure 6

a PWY-7221: guanosine ribonucleotides de novo biosynthesis



b

Bacteroides vulgatus
 $r = 0.2104$, $p = 0.0123$



c

Alistipes putredinis
 $r = -0.1848$, $p = 0.0282$

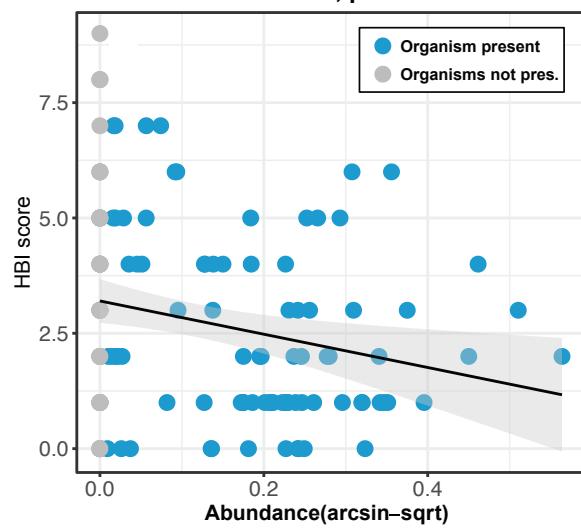


Figure S6: (a) Pathways with the greatest contributonal alpha diversity. In addition to the scaled versions (shown on the right), the corresponding plots showing the relative abundance of the pathway in the samples ($n=78$) are displayed on the left. (b) and (c) Associations of *B. vulgatus* and *A. putredinis* taxonomic abundances with HBI score (a disease activity index) across all 300 metagenomic samples. r and p -values are Spearman correlations, regression line is a linear fit.

Supplementary Table 1

Species	Perimeter of triangle (log10)
<i>Ruminococcus gnavus</i>	6.48
<i>Bacteroides faecis</i>	5.85
<i>Clostridium symbiosum</i>	4.66
<i>Bacteroides fragilis</i>	4.6
<i>Clostridium clostridioforme</i>	4.43
<i>Roseburia intestinalis</i>	3.71
<i>Eubacterium hallii</i>	3.65
<i>Roseburia hominis</i>	3.57
<i>Alistipes onderdonkii</i>	3.43
<i>Clostridium hathewayi</i>	3.37
<i>Bacteroides xylinisolvans</i>	3.32
<i>Ruminococcus obeum</i>	3.02
<i>Clostridium bolteae</i>	2.83
<i>Parabacteroides distasonis</i>	2.73
<i>Bacteroides dorei</i>	2.56
<i>Bilophila wadsworthia</i>	2.31
<i>Eubacterium rectale</i>	2.3
<i>Dorea longicatena</i>	2.22
<i>Ruminococcus torques</i>	2.18
<i>Akkermansia muciniphila</i>	2.14
<i>Alistipes finegoldii</i>	1.99
<i>Ruminococcaceae bacterium D16</i>	1.98
<i>Haemophilus parainfluenzae</i>	1.82
<i>Burkholderiales bacterium 1 1 47</i>	1.74
<i>Eubacterium eligens</i>	1.69
<i>Sutterella wadsworthensis</i>	1.69
<i>Parasutterella excrementihominis</i>	1.68
<i>Lachnospiraceae bacterium 5 1 63FAA</i>	1.61
<i>Flavonifractor plautii</i>	1.61
<i>Roseburia inulinivorans</i>	1.54
<i>Escherichia coli</i>	1.51
<i>Lachnospiraceae bacterium 7 1 58FAA</i>	1.44
<i>Anaerostipes hadrus</i>	1.31
<i>Faecalibacterium prausnitzii</i>	1.24
<i>Bacteroides caccae</i>	1.21
<i>Bacteroides vulgatus</i>	1.19
<i>Bacteroides stercoris</i>	1.15
<i>Bacteroides ovatus</i>	1.1
<i>Bacteroides thetaiotaomicron</i>	1.06
<i>Parabacteroides merdae</i>	0.89
<i>Bacteroides uniformis</i>	0.82
<i>Alistipes putredinis</i>	0.65

Table S1: Complete list of all species and their corresponding triangle perimeter representing their metagenomic and metatranscriptomic activity within each disease group. Values are averaged over samples from the same individual and then across individuals. The species with the 7 largest and 15 smallest changes are displayed in Figure 2D+E.